## CS-E4850 Computer Vision

Exam 24th of May 2017, Lecturer: Juho Kannala

There are plenty of questions, answer as many as you can in the available time. The number of points awarded from different parts is shown in parenthesis in the end of each question. The maximum score from the whole exam is 42 points.

You will need pen and paper, and also calculator is allowed but is not necessary.

1. Explain briefly the following concepts (e.g. what does the concept mean, what are its key properties, and how it is utilised in computer vision):
   (a) Separable filter (2 p)
   (b) Scale invariant feature transform (SIFT) (2 p)
   (c) Camera calibration (2 p)
   (d) Structure from motion (2 p)
   (e) Convolutional neural network (2 p)
   (f) Viola-Jones face detector (2 p)

2. Model fitting using RANSAC algorithm
   (a) Describe the main stages of the RANSAC algorithm in the general case. (2 p)
   (b) Mention at least two examples of models that can be fitted using RANSAC. Describe how the models are used in computer vision and what is the size of the minimal subset of data points required for fitting in each case. (2 p)
   (c) Describe how RANSAC can be used for panoramic image stitching. Why is RANSAC needed and what is the model fitted in this case? (2 p)

3. Large-scale object instance recognition
   (a) Describe the bag-of-visual-words image representation technique and its pros and cons for object instance recognition. (2 p)
   (b) Describe what is *inverted index* and how it can be used to improve efficiency of object instance recognition from large image databases? (1 p)
   (c) Explain the concept *term frequency - inverse document frequency* (tf-idf) weighting and its purpose. (1 p)
   (d) Explain what is the precision-recall curve (that is often used for evaluating image retrieval systems). (2 p)

4. Feature tracking
   (a) Describe the main elements of the Shi-Tomasi feature tracker (i.e. how the features are selected and tracked, and tracks terminated or added). (2 p)
   (b) Describe the Lucas-Kanade method for estimating the displacement of an image patch. What kind of equations need to be solved and what is the brightness constancy constraint in this context? (2 p)
   (c) What are the benefits of Lucas-Kanade method when compared to simple template matching? (2 p)

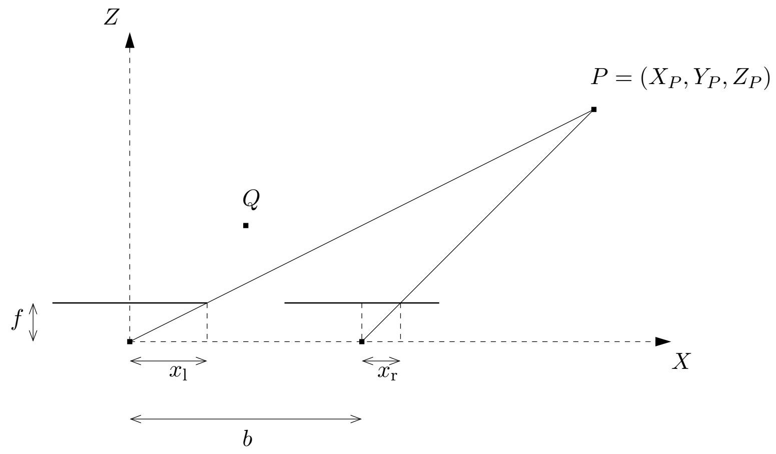


Figure 1: Top view of a stereo pair where two pinhole cameras are placed side by side.

5. Epipolar geometry and stereo

   (a) Figure 1 presents a stereo system with two parallel pinhole cameras separated by a baseline $b$ so that the centers of the cameras are $\mathbf{c}_l = (0, 0, 0)$ and $\mathbf{c}_r = (b, 0, 0)$. Both cameras have the same focal length $f$. The point $P$ is located in front of the cameras and its disparity $d$ is the distance between corresponding image points, i.e., $d = |x_l - x_r|$. Assume that $d = 1$ $cm$, $b = 6$ $cm$ and $f = 1$ $cm$. Compute $Z_P$. (2 p)

   (b) Let's denote the camera projection matrices of two cameras by $\mathbf{P} = \begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}$ and $\mathbf{P}' = \begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}$, where $\mathbf{R}$ is a rotation matrix and $\mathbf{t} = (t_1, t_2, t_3)^\top$ describes the translation between the cameras. Show that the epipolar constraint for corresponding image points $\mathbf{x}$ and $\mathbf{x}'$ can be written in the form $\mathbf{x}'^\top \mathbf{E}\mathbf{x} = 0$, where matrix $\mathbf{E}$ is the essential matrix $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$. (2 p)

   (c) In the configuration illustrated in Figure 1 the camera matrices are $\mathbf{P} = \begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}$ and $\mathbf{P}' = \begin{bmatrix}\mathbf{I} & \mathbf{t}\end{bmatrix}$, where $\mathbf{I}$ is the identity matrix and $\mathbf{t} = (-6, 0, 0)^\top$. The point $Q$ has coordinates $(3, 0, 3)$. Compute the image of $Q$ on the image plane of the camera on the left and the corresponding epipolar line on the image plane of the camera on the right. (Hint: The epipolar line is computed using the essential matrix.) (2 p)

6. Geometric 2D transformations

   (a) Using homogeneous coordinates, write the matrix form of the following 2D transformations: translation, similarity (rotation+scaling+translation), affine and homography. How many degrees of freedom does each transformation have? How many point correspondences are needed to estimate each? (3 p)

   (b) A rectangle with corners $A = (-1, 1)$, $B = (1, 1)$, $C = (1, -1)$, $D = (-1, -1)$ is transformed by a transformation so that the new corners are $A' = (1, 3)$, $B' = (3, 3)$, $C' = (-2, 1)$, $D' = (-6, 1)$, respectively. An affine transformation does not explain the observations perfectly, but there is reason to believe that the transformation is affine and there is noise in the observations. Write down the equations to solve the transformation using the least squares method.
   **Note:** You don't actually have to solve the transformation. (3 p)